# CHAPTER 4

# ASSESSMENT OF SKILL TRANSFER

## 4.1 Aim

In this final part of the thesis, investigations have been made not only about pupils' ability to retain the skills, but also his/her ability to transfer the skills to a situation different from that encountered during learning sessions conducted earlier.
During the school period, a pupil's primary objective for learning a skill, retaining and reproducing it in an assessment session is often to gain a good grade. However, in real life the purpose of learning process is much more than passing an examination. After leaving school, a pupil would be expected to be involved in his/her working life, where he/she encounters problems different from those met in a school laboratory. In that situation the person would have to recall the skill learnt (or knowledge gained) and transfer it to a new situation to solve the recently encountered problem. Assessment of skill transfer ability of pupil is hence an important part of the practical examination process and this part of the thesis has been used to study it.

For this purpose, a group of pupils has been taught certain manipulative skills. After a certain period of time, they were assessed for executing manipulative skills, which are similar but not the same, i.e. two different skills having certain features in common. A control group (untaught group) have also been assessed to use their performance as a standard for comparing the results of taught group.

## 4.2 Development of the study materials --- preliminary runs

### (i) Method

No trial run was carried out either for the teaching sessions or the testing sessions used to assess the pupils' ability to retain the skills taught earlier and then to transfer it into a new situation. However, experience gained during the work involved in chapters 2 and 3 was taken into consideration for organising the test session in this part of the investigation. In addition to this, great care was taken to make the worksheets simple and easy to follow, and also to ensure an adequate supply of chemicals and apparatus. For teaching sessions, two experiments were chosen, one experiment involves measurement of volume by using a measuring cylinder while the other experiment concerns heating solids in a test-tube over Bunsen burner. Two different experiments were chosen for testing sessions. This time one experiment involves measuring volume using a burette and the other involves heating liquid in a test-tube over a Bunsen burner.

### 4.3 Final run

(i) Method

Two sets of third form pupils (13-14 years old), each comprising of about 24 boys, from a local independent school (Norwich School, The Close, Norwich) took part in this investigation. At the beginning of the spring term they were given lessons by four experienced science teachers (two teachers per set) on a number of scientific skills involved in two experiments: (i) Heat of reaction between acid and alkali (experiment 4.1), and (ii) Heating solid samples in test tubes (experiment 4.2) [see work-sheets 4.1 and 4.2 in Appendix 3.1].

For these teaching sessions, each set of pupils was randomly divided into three groups: Group I was given a lesson on the measurement of volume of acid/alkali involved in experiment 4.1, Group II was given lessons on

heating solids in test-tubes, and Group III was kept as the Control group. The pupils in Group III were shown a video film, unrelated to either two of the experiments mentioned above.
For this teaching activity, two laboratories in Norwich School were used. One was equipped with materials needed for the volume measurement experiment, using a measuring cylinder and pupils belonging to Group I from both the sets were taught. The other laboratory was equipped with materials needed for the heating of solids in test tubes where the pupils belonging to Group II from both the sets were taught. There were two teachers in each laboratory and the pupils were given lessons on performing several manipulative skills accurately over a period of one hour. Chemicals and apparatus used in the teaching sessions (experiment 4.1 and 4.2) are enlisted in Appendix 3.2.

The pupils in Group I (measurement of volume group) were taught how to measure accurately a definite volume of liquid by using a measuring cylinder. It was stressed that the lower meniscus must touch the required volume mark on the cylinder. The pupils in Group II (heating group) were taught how to follow the necessary steps involved in heating a solid sample in a test-tube. In heating solids, the pupils were expected to do the following:- (a) place the test-tube holder near the top of the tube, (b) hold the tube at about $45^{0}$ angle and not pointing towards any person, (c) place the end of the tube in the hot zone of the flame, and (d) use the blue flame. These four steps have been thought to be important in influencing the results of the thermal treatment of solids.

After a period of about one and a half months, the pupils were tested. Now, all the pupils in each set were given experimental work and each pupil had to do two experiments. The first experiment (experiment 4.3) involved the measurement of a certain volume of liquid (water and alcohol) using a burette and to weigh it in order to determine its density (see work-sheet 4.3 in Appendix 3.1). The second experiment (experiment 4.4) dealt with the heating of a liquid in test-tube (see work-sheet 4.4. in Appendix 3.1). For this pupils were asked to heat a series of sugar solutions, mixed with Benedict's solution and identify whether they were reducing or non-reducing sugars.

For this testing session two science laboratories (the same as those used in the teaching session) in Norwich School were used. One laboratory had facilities for carrying out the experiments involving measurement of volume by using a burette, while the other laboratory had facilities for the experiment involving heating solutions in test tubes using a Bunsen burner.

Two sets of pupils, each containing samples from Group I, II and III took part, and carried out both the experiments which lasted for about an hour. In the first half of the session one set did volume measurement experiments, while the other set did heating experiments. Then in the second half of the session, the two sets swapped laboratories as well as experiments. The lists of chemicals and apparatus used in experiment 4.3 and 4.4 are given in Appendix 3.2.

The work-stations were numbered in the same way as described for the testing session in the previous chapters. The pupils were randomised and given number labels at the entrance of the laboratory at the beginning of the first part of the session, and this numbering remained valid for the

second half of the session, even though the whole set of pupils swapped laboratories after the first half of the session.
The performance of the pupils in these two experiments, which took about an hour, was judged by observation (Method -T) by experienced teachers who had no knowledge of the teaching sessions. A panel of 3 judges assessed the measurement of volume skill in experiment 4.3 for both the sets, and another panel of 3 judges assessed the heating skills in experiment 4.4 for both the sets. Before the start of the experiments, the pupils in each set were randomly allocated to various work-stations.

The judges were given clear instructions regarding their assessment scheme. For the heating experiments (experiment 4.4), 4 important skills, mentioned earlier in the description of the teaching sessions, were checked. Each of the judges assessed all the pupils once (4 skills). As there were 3 judges, each pupil obtained a maximum of 3 marks for performing each heating skill correctly. In other words, each pupil was assessed thrice altogether and obtained a maximum of 12 marks (3 x 4) for correct execution of all the 4 skills. During the measurement of volume experiment (experiment 4.3), one judge (in the panel of three) was able to assess only 8 pupils (one third of the set) stationed at one bench, during the half hour taken by the set to complete all the experimental work on 5 samples. In other words, each pupil was assessed only once and obtained 1 mark for measuring accurately a certain volume of liquid (within an error limit of ± 0.1 ml) by using a burette. For this purpose, each of the three judges stayed near a set of burettes containing a series of liquids to be used by 8 pupils at a bench and noted the burette readings before and after a pupil drained out the liquid. There were three rows of burettes, attended by three judges, to be used by pupils from three benches. These burettes were filled from time to time.

In addition to the assessment by observation of the skill-performance (Method -T), the final results of the volume measurement experiment, which were the density values of a series of liquids, were also assessed. This provided a basis for assessing the skill-performance from another angle and is called Method -P. Experiments were carried out by me (5 runs per sample) to find out the exact density values for all the samples and these data (with the standard deviation as the error limit) were used as standards for assessing the results reported by the pupils, (see Appendix 3.3). As there were 5 samples, each pupil was awarded a maximum of 5 marks for reporting correct density data.

## 4.3 Final Run

### (ii) Results and Discussion

The information available in the judges' assessment sheets (see Appendix 3.4 ) has been collated and given in Table 4.1 and 4.2. This provided the achievement data for pupils (assessed by Method -T) belonging to three groups in each of the two sessions in carrying out (1) measurement of volume using burettes involved in experiment 4.3 (see work-sheet 4.3 in appendix 3.1) and (2) heating solutions in test-tubes using Bunsen burner in experiment 4.4 (see work-sheet 4.4 in appendix 3.1). The achievement of these pupils for carrying out volume measurement based on the assessment (by Method -P) of their final results (i.e. accuracy of the density data) of experiment 4.3 has also been determined. Finally these data in Tables 4.1 and 4.2 have been analysed for the mean values of achievement, together with standard deviation, of pupils belonging to each of the three groups and are given in Table 4.3.

In experiment 4.3, the group who had lessons on volume measurement using a measuring cylinder have done better than the other two groups even though this time the volume measurement was done by using a burette. The other two groups seem to have done quite well, although they had no lessons on the skill of measurement of volume.

The achievement data obtained by assessment Method -P (i.e. marking the reported density values) of experiment 4.3 also suggest that the group who had lessons on volume measurement skill performed better than the other two groups. However, each group's achievement is so high that one can conclude that the teaching of volume measurement skill has probably no significant effect on the final results (i.e. density values) determined in this way. Pearson -r between the achievement when assessed by Method -T and that when assessed by Method -P for (i) volume measurement group, (ii) heating group, and (iii) control group have been calculated and are given in Table 4.4.

As Pearson -r value for each of these 3 groups lies within +0.5 and -0.5, it is unlikely that there is a strong correlation between the two variables. However, it is worthwhile looking at the critical values of "r" given in the statistical table and finding out the statistical significance of the calulated Pearson -r values. All the 3 "r" values seem to be lower than the critical "r" values for df = 12, 12, 11 respectively even at $p = 0.05$ indicating that % probability of correlation by chance factor is higher than 5%. Thus the Null Hypothesis can be accepted at this level stating that the existing correlation is due to chance factor and there is no significant correlation between the two sets of data. However, although the experimental results do not provide any support for the hypothesis that improvement in pupils' ability to measure liquid volume by burette accurately will make an improvement in the accuracy of the reported density data of this liquid, the hypothesis cannot be rejected completely. This is because the total number of samples in each group taking part in the experiment was relatively small compared with a Nation-wide situation having a large number of samples which may provide a more meaningful correlation supportive of the experimental hypothesis. Nevertheless, the design of experiments used in this study was sound statistically, as each method group contained almost equal numbers of randomised samples of both sexes drawn from a pool of 40 pupils (2 sets) which could be considered as a representative of the pupils of the same age group in the country. Figures 4.1, 4.2 and 4.3 show the plots of the achievement data while assessed by Method -T against that while assessed by Method -P for measuring volume using a burette by each of the three groups of pupils. No clear trend or pattern can be seen in any of these figures which supports the prediction from the Pearson-r value calculation that no significant correlation exists between the achievement data obtained from two different mode of assessment of the volume measurement skill of any of the three groups.

Results of experiment 4.4 indicate that the group which had lessons on heating solids in test tubes performed better in heating solutions in test tubes than the other two groups. The heating group's better performance is clearly visible in all the four skills involved in experiment 4.4. However, except for the task concerned with the skill of position of the test tube holder, the mean value of achievement is considerably higher for all the three groups. It may be due to the fact that teaching had very little impact on the performance of the heating group.

Several 3 X 2 ANOVA based on randomised groups were done. They were as follows:

i) In experiment 4.3

| | |
|---|---|
| Assessment<br>of volume measurement ) | 3 method group(A)<br>X 2 sets (B) |

(ii) In experiment 4.3

| | |
|---|---|
| Assessment<br>of the final results<br>(density data) ) | 3 method group(A)<br>X 2 sets (B) |

(iii) In experiment 4.4

Skill (a):- 3 method groups (A) X 2 sets (B)

(iv) In experiment 4.4.

Skill (b):- 3 method groups (A) X 2 sets (B)

(v) In experiment 4.4.

Skill (c):- 3 method groups (A) X 2 sets (B)

(vi) In experiment 4.4.

Skill (d):- 3 method groups (A) X 2 sets (B)

---

Information derived from the ANOVA is in Table 4.5.

None of the $F_A$, $F_B$ or $F_{AXB}$ seem to be significant even at $p < 0.05$ level.

From the analysis of variance, it appears that the teaching of a skill to a group had no significant effect in the achievement, we have found that for a few skills the taught group was slightly better than the other two groups and for the other skills all the three group's achievements were very similar. However, this does not necessarily mean that higher achievement data for the taught group observed in the assessment of a few skills has no significance at all. The lack of statistical significance of $F_A$ at $p < 0.05$ level could be due to high standard deviation value of the data. Nevertheless, $F_A$ could be meaningful at a level slightly higher than $p = 0.05$ and thus it is worthwhile to recognise the higher achievement (although small) of the taught group compared to the untaught group in extending the skill to a new situation.. The ANOVA also suggest that the results of assessment has not been significantly influenced by having the subjects in 2 sets and assessing them in 2 separate sessions.

The lack of correlation between the achievement data obtained from two methods of assessment raises questions regarding the assumption made earlier that "such a comparison can be used as a measure of validity of the method of assessment". Now one may ask if both the methods are invalid or either of the methods is invalid. If the latter case is correct, which one is invalid?

To address these questions, I would suggest that perhaps one of these methods cannot be used to validate the other. The basic reason is, as it appears from the findings of this research work, that we are not measuring the same thing in two methods of assessment. We are measuring two different aspects of the same experiment.

To clarify this statement I would state that in the Method -T of assessment, we are measuring a few specific psychomotor skills, but in Method -P we are assessing the outcome of the extended task which involves the effect of skills already assessed by Method -T plus some other skills which are required to complete the whole experiment. So the picture produced by the achievement data obtained from Method -P does not give a mirror image of that obtained from the assessment Method -T.

This suggests that we cannot substitute the assessment Method -T by Method -P, even though the latter method is more convenient for a teacher, as it can be carried out at a later stage, i.e. outside the busy lesson time. However, assessment by Method -T is essential as it allows us to observe in-situ how a pupil performs each manipulative skill. Assessment by Method -P, on the other hand, deals with the cumulative effect of a series of manipulative skills and cognitive skills which is submitted as the final result of the whole experiment.

Thus one of these methods cannot be used to measure the validity of the other. Assessment of two different aspects of an experiment by two methods is likely to be the reason for not obtaining a significant correlation between the achievement data of pupils assessed by these two methods of assessment.

To validate one method by another, it is vital to ensure that both the methods are measuring the same thing. For example, we can measure the thickness of a piece of glass by using a ruler and we can find the validity of this method by measuring again the same thickness by a more sophisticated technique, viz. slide callipers or spherometer.

In the work described in this thesis, the situation seems to be different. The two methods, T and P are not assessing the same aspects of an experiment: instead they are assessing two different aspects of the same experiment.

Let us consider the experiment 2.1 in Chapter 2. Here pupils measured the volume of liquid by a measuring cylinder, and poured it out into a flask containing Mg. These two skills are assessed by Method -T. One may expect that if these two skills are performed correctly, the time needed for Mg ribbon to dissolve in acid completely (basis of the assessment by Method -P) would be correct too and pupils are expected to obtain full marks while assessed from two directions (Method -T and -P) providing a significant correlation between the achievement data obtained from two methods of assessment. But a careful analysis of the steps involved in the work assessed by two methods would reveal that the assessment by Method -P deals with a few more aspects of the pupils' ability of performing practical skills in addition to those already assessed by Method -T. These extra skills will have influence on the results and consequently the value reported as the final result of the whole experiment is not representing exactly the same skills of the pupil as assessed by Method -T. Similar arguments can be put forward regarding the assessment of two different aspects of experiment 4.3 in Chapter 4 by two different methods. This explains why we are not obtaining a significant correlation between the pupils' achievement data obtained from two methods of assesment used in each of these two experiments.

To validate either Method -T or Method -P of assessment, it would be necessary to design (formulate) another procedure which measures exactly the same thing.

To illustrate this idea, I would say that in order to validate assessment by

Method -T of pupils' ability to measure volume of liquid by a measuring cylinder, one could repeat the experiment by using a burette and still assess the skill performance by Method -T. Similarly to validate the assessment of the final results of the whole experiment submitted by the pupil as a written report (i.e. Method -P), one could devise another method for completing the subsequent steps in experiment 2.1 (Chapter 2) or 4.3 (Chapter 4) by some other sophisticated means. Thus two ways of carrying out the same aspects of an experiment and both assessed by Method -P become worth comparing.

Table 4.1

Total marks obtained by the pupils in set 1 for carrying out volume measurement by burette in experiment 4.3 and heating liquids in test tube experiment 4.4 (*assessed by method -P) (see page 79 and 80 for explanation of max. marks)

| Pupil No. in (Set 1) | | Experiment 4.3 (session 2) max marks | | Experiment 4.4 (session 1) | | | |
|---|---|---|---|---|---|---|---|
| | | = 1 | = 5* | Position of test tube holder max marks = 3 | Angle of test tube max marks = 3 | Zone of flame max marks = 3 | Colour of flame max marks = 3 |
| 1 | H | 1 | 3 | 3 | 3 | 3 | 3 |
| 2 | C | 1 | 2 | 0 | 3 | 3 | 3 |
| 3 | V | 1 | 3 | 0 | 1 | 2 | 3 |
| 4 | H | 1 | 2 | 3 | 3 | 3 | 3 |
| 5 | - | - | - | - | - | - | - |
| 6 | V | 1 | 3 | 0 | 3 | 3 | 3 |
| 7 | C | 0 | 2 | 0 | 2 | 3 | 3 |
| 8 | V | 1 | 1 | 1 | 3 | 2 | 2 |
| 9 | V | 0 | 3 | 3 | 3 | 3 | 2 |
| 10 | C | 0 | 0 | 0 | 1 | 1 | 2 |
| 11 | C | 0 | 2 | 3 | 3 | 2 | 1 |
| 12 | - | - | - | - | - | - | - |
| 13 | H | 0 | 3 | 3 | 3 | 3 | 3 |
| 14 | H | 1 | 1 | 3 | 2 | 3 | 3 |
| 15 | V | 0 | 4 | 3 | 2 | 0 | 1 |
| 16 | C | 1 | 3 | 1 | 3 | 2 | 2 |
| 17 | V | 1 | 4 | 3 | 3 | 3 | 3 |
| 18 | H | 0 | 4 | 1 | 3 | 2 | 3 |
| 19 | - | - | - | - | - | - | - |
| 20 | H | 1 | 4 | 2 | 3 | 3 | 3 |
| 21 | C | 1 | 1 | 3 | 3 | 3 | 3 |
| 22 | V | 0 | 3 | 3 | 3 | 2 | 3 |
| 23 | C | 0 | 4 | 1 | 3 | 0 | 3 |
| 24 | - | - | - | - | - | - | - |
| 25 | H | 0 | 3 | 3 | 3 | 3 | 3 |

V/H/C in column one indicate the group (volume measurement/heating solids/control) they belonged to during the teaching sessions earlier.

Table 4.2

Total marks obtained by the pupils in set 2 for carrying out volume measurement by burette in experiment 4.3 and heating liquids in test tubes in experiment 4.4 (*assessed by method -P) (see page 79 and 80 for explanation of max marks)

| Pupil No. in (Set-2) | Experiment 4.3 (session 1) max mark | | Experiment 4.4 (session 2) | | | |
|---|---|---|---|---|---|---|
| | 1 | 5* | Position of test tube holder max mark = 3 | Angle of test tube max mark = 3 | Zone of flame max mark = 3 | Colour of flame max mark = 3 |
| 1 C | 1 | 3 | 2 | 2 | 3 | 3 |
| 2 V | 1 | 2 | 3 | 2 | 3 | 3 |
| 3 V | 1 | 4 | 3 | 2 | 2 | 3 |
| 4 V | 1 | 2 | 0 | 3 | 3 | 3 |
| 5 V | 1 | 3 | 1 | 3 | 2 | 3 |
| 6 C | 1 | 4 | 2 | 2 | 3 | 2 |
| 7 H | 1 | 1 | 3 | 3 | 3 | 3 |
| 8 H | 1 | 3 | 3 | 3 | 3 | 3 |
| 9 H | 1 | 2 | 0 | 2 | 3 | 3 |
| 10 V | 1 | 3 | 3 | 1 | 3 | 3 |
| 11 C | 1 | 2 | 0 | 2 | 3 | 3 |
| 12 H | 1 | 2 | 1 | 3 | 3 | 3 |
| 13 C | 1 | 2 | 2 | 3 | 3 | 3 |
| 14 C | 0 | 4 | 2 | 3 | 2 | 3 |
| 15 – | – | – | – | – | – | – |
| 16 C | 0 | 4 | 0 | 2 | 3 | 3 |
| 17 H | 0 | 3 | 3 | 3 | 3 | 2 |
| 18 H | 1 | 2 | 0 | 2 | 3 | 3 |
| 19 C | 0 | 1 | 0 | 2 | 3 | 3 |
| 20 H | 1 | 3 | 3 | 3 | 0 | 3 |
| 21 H | 0 | 2 | 1 | 2 | 2 | 2 |
| 22 V | 0 | 3 | 3 | 3 | 3 | 3 |
| 23 V | 1 | 2 | 3 | 2 | 3 | 3 |
| 24 V | 1 | 4 | 0 | 3 | 1 | 3 |
| 25 – | – | – | – | – | – | – |

V/H/C in column one indicate the group (volume measurement/heating solid/control) they belonged to during the teaching sessions earlier.

## Table 4.3

Mean values of the achievement of different groups of pupils in carrying out several practical skills in two experiments: 4.3 and 4.4.

| Expt. | Practical skill assessed | Measurement of volume gr. | | Heating of solids gr. | | Control Group | | Max. value of x |
|---|---|---|---|---|---|---|---|---|
| | | x | s.d. | x | s.d. | x | s.d. | |
| 4.3 | Measurement of volume | 0.73 | ±0.44 | 0.67 | ±0.47 | 0.50 | ±0.50 | 1 |
| | * | 2.93 | ±0.85 | 2.53 | ±0.88 | 2.43 | ±1.24 | 5 |
| 4.4 | (a) Position of the test tube holder | 1.93 | ±1.34 | 2.13 | ±1.15 | 1.14 | ±1.12 | 3 |
| | (b) Angle of the test tube | 2.47 | ±0.72 | 2.73 | ±0.44 | 2.43 | ±0.62 | 3 |
| | (c) Zone of flame | 2.33 | ±0.87 | 2.66 | ±0.79 | 2.43 | ±0.90 | 3 |
| | (d) Colour of flame | 2.73 | ±0.57 | 2.87 | ±0.34 | 2.64 | ±0.61 | 3 |

(x = mean value, s.d. = standard deviation)

* Results of the skill assessed by Method -P(marking the final outcome of the experiment, i.e. density data).
[The rest of the data in the table are results of assessment by method -T (on-the-spot observation by the judges)]
(See page 79 and 80 for explanation of maximum value of x).

# TABLE 4.4

Pearson-r (Pearson product moment correlation coefficient) between the achievements of pupils after being assessed by Method -T and that after being assessed by Method -P for their volume measurement skill in experiment 4.3.

| Volume measurement Group | Heating Group | Control Group |
|---|---|---|
| -0.223 | -0.373 | 0.133 |

## Table 4.5

Summary tables for 3X2 completely randomised ANOVA of the results of studies on pupils' skill transfer abilities. (see experiments 4.3 and 4.4 in Chapter 4).

(a) Experiment 4.3 Measurement of volume of some liquids by burette (assessed by the judges on the spot, i.e. Method -T)

| Source of variance | SS | df | MS | F (obsv.) | F (crit) at p=0.05 | Significance of F (obsv.) |
|---|---|---|---|---|---|---|
| A (method group) | 0.429 | 2 | 0.2145 | 0.9003 | 3.32 | n.s. ($p < 0.05$) |
| B (class/set) | 0.595 | 1 | 0.595 | 2.499 | 4.17 | n.s. ($p < 0.05$) |
| A × B | 0.0479 | 2 | 0.024 | 0.1005 | 3.32 | n.s. ($p < 0.05$) |
| Error | 8.571 | 36 | 0.238 | - | - | - |
| Total | 9.643 | 41 | - | - | - | - |

(b) Experiment 4.3 Measuring volume by burette (assessed on the basis of reported density data, i.e. Method -P)

| Source of variance | SS | df | MS | F (obsv.) | F (crit) at p=0.05 | Significance of F (obsv.) |
|---|---|---|---|---|---|---|
| A (method group) | 1.334 | 2 | 0.667 | 0.623 | 3.32 | n.s. ($p < 0.05$) |
| B (class/set) | 0.00024 | 1 | 0.00024 | 0.000224 | 4.17 | n.s. ($p < 0.05$) |
| A × B | 3.999 | 2 | 1.9995 | 1.867 | 3.32 | n.s. ($p < 0.05$) |
| Error | 38.572 | 36 | 1.0714 | - | - | - |
| Total | 43.905 | 41 | - | - | - | - |

/continued

## Table 4.5

(c) Experiment 4.4 Heating of solutions in test tube.
(combined four skills, assessed by Method -T)

| Source of variance | SS | df | MS | F (obsv.) | F (crit) at p=0.05 | Significance of F (obsv.) |
|---|---|---|---|---|---|---|
| A (method group) | 28.04 | 2 | 14.02 | 2.82 | 3.32 | n.s. ($p < 0.05$) |
| B (class/set) | 0.85 | 1 | 0.425 | 0.17 | 4.17 | n.s. ($p < 0.05$) |
| A × B | 13.01 | 2 | 6.505 | 1.31 | 3.32 | n.s. ($p < 0.05$) |
| Error | 179.00 | 36 | 4.972 | - | - | - |
| Total | 220.90 | 41 | - | - | - | - |

(d) Experiment 4.4 Heating of solutions in test tube,
skill tested: Position of the test tube holder
(assessed by Method -T)

| Source of variance | SS | df | MS | F (obsv.) | F (crit) at p=0.05 | Significance of F (obsv.) |
|---|---|---|---|---|---|---|
| A (method group) | 9.48 | 2 | 4.74 | 3.02 | 3.32 | n.s. ($p < 0.05$) |
| B (class/set) | 0.10 | 1 | 0.10 | 0.064 | 4.17 | n.s. ($p < 0.05$) |
| A × B | 2.33 | 2 | 1.165 | 0.74 | 3.32 | n.s. ($p < 0.05$) |
| Error | 56.57 | 36 | 1.571 | - | - | - |
| Total | 68.48 | 41 | - | - | - | - |

(e) Experiment 4.4 Heating of solutions in test tube.
skill tested: Angle of the test tube
(assessed by Method -T)

| Sourcr of Variance | SS | df | MS | F (obsv.) | F (crit) at p=0.05 | Significance of F (obsv.) |
|---|---|---|---|---|---|---|
| A (method group) | 1.186 | 2 | 0.593 | 1.467 | 3.32 | n.s. ($p < 0.05$) |
| B (class/set) | 0.59 | 1 | 0.59 | 1.460 | 4.17 | n.s. ($p < 0.05$) |
| A × B | 0.06 | 2 | 0.03 | 0.074 | 3.32 | n.s. ($p < 0.05$) |
| Error | 14.564 | 36 | 0.405 | - | - | - |
| Total | 16.4 | 41 | - | - | - | - |

/continued

## Table 4.5

(f) Experiment 4.4 Heating of solutions in test tube.
skill tested: Zone of flame
(assessed by Method -T)

| Source of variance | SS | df | MS | F (obsv.) | F (crit) at p=0.05 | Significance of F (obsv.) |
|---|---|---|---|---|---|---|
| A (method group) | 0.76 | 2 | 0.38 | 0.54 | 3.32 | n.s. ($p < 0.05$) |
| B (class/set) | -0.47 | 1 | -0.47 | -0.66 | 4.17 | n.s. ($p < 0.05$) |
| $A \times B$ | 4.47 | 2 | 2.235 | 3.14 | 3.32 | n.s. ($p < 0.05$) |
| Error | 25.72 | 36 | 0.714 | - | - | - |
| Total | 30.48 | 41 | - | - | - | - |

(g) Experiment 4.4 Heating of solutions in test tube.
skill tested: Colour of flame
(assessed by Method -T)

| Source of variance | SS | df | MS | F (obsv.) | F (crit) at p=0.05 | F (obsv.) |
|---|---|---|---|---|---|---|
| A (method group) | 0.62 | 2 | 0.31 | 1.24 | 3.32 | n.s. ($p < 0.05$) |
| B (class/set) | 0.86 | 1 | 0.86 | 3.44 | 4.17 | n.s. ($p < 0.05$) |
| $A \times B$ | 1.00 | 2 | 0.50 | 2.00 | 3.32 | n.s. ($p < 0.05$) |
| Error | 9.14 | 36 | 0.254 | - | - | - |
| Total | 11.62 | 41 | - | - | - | - |

In Tables 4.5 (a) - (g):

| | |
|---|---|
| F (obsv.) | denotes the value of F ratio obtained from the experimental results. |
| F (crit) | denotes the critical value of F ratio at a given level of significance (i.e. p value) and at the appropriate df values (see reference 50 for the relevant statistical tables). |
| n.s. | means "not significant". This refers to the fact that the value of observed F ratio is less than the critical value of F at p=0.05 and thus F (observed) is not significant at $p < 0.05$ level. |

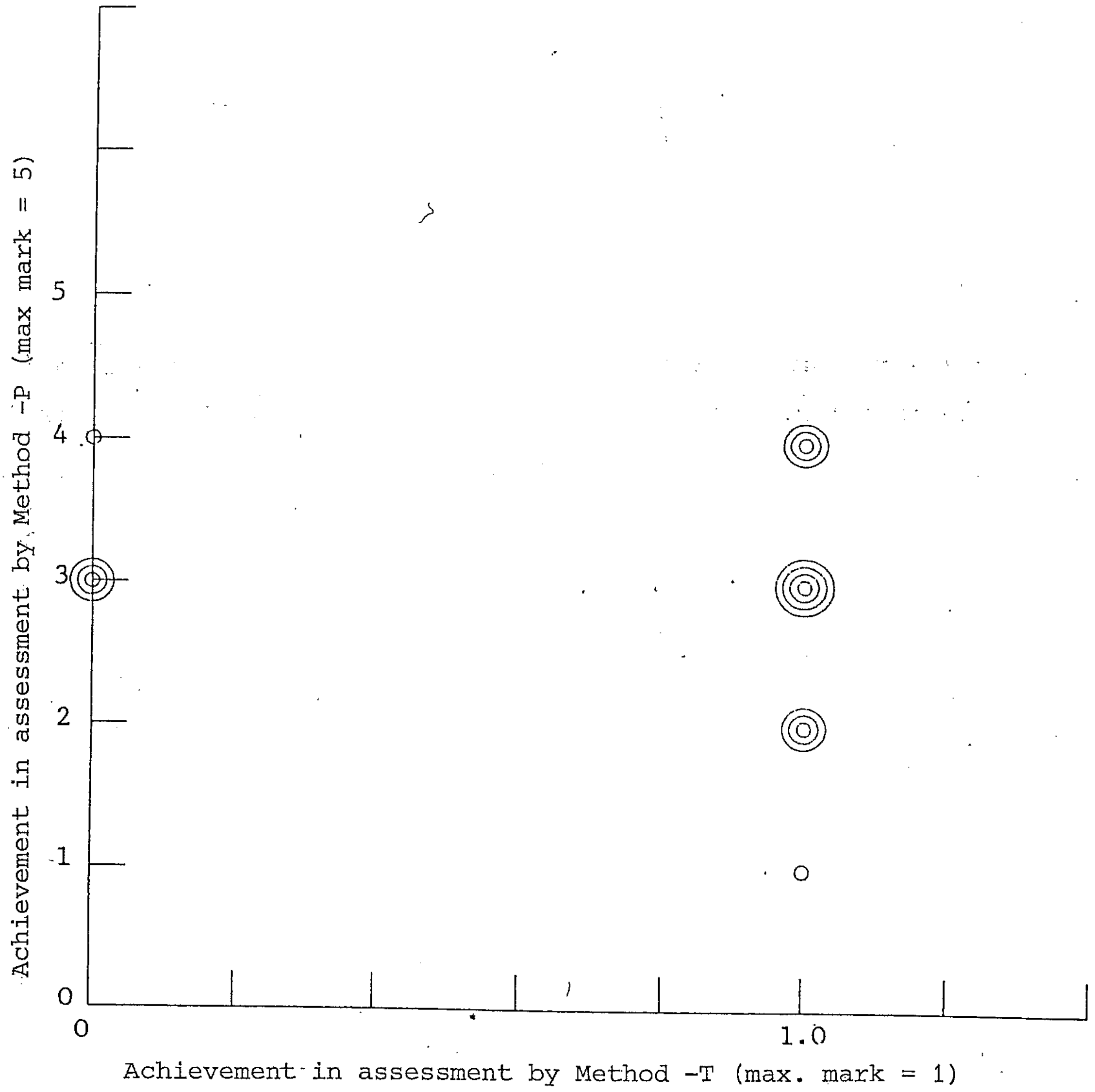

Fig. 4.1 Plots of achievement data obtained from assessment by Method -T for measuring volume using burette by pupils of the "Volume measurement" group (those who got lessons on measuring volume by a measuring cylinder) participating in both the sessions against their achievement data obtained from assessment by Method -P for carrying out the whole experiment 4.3.

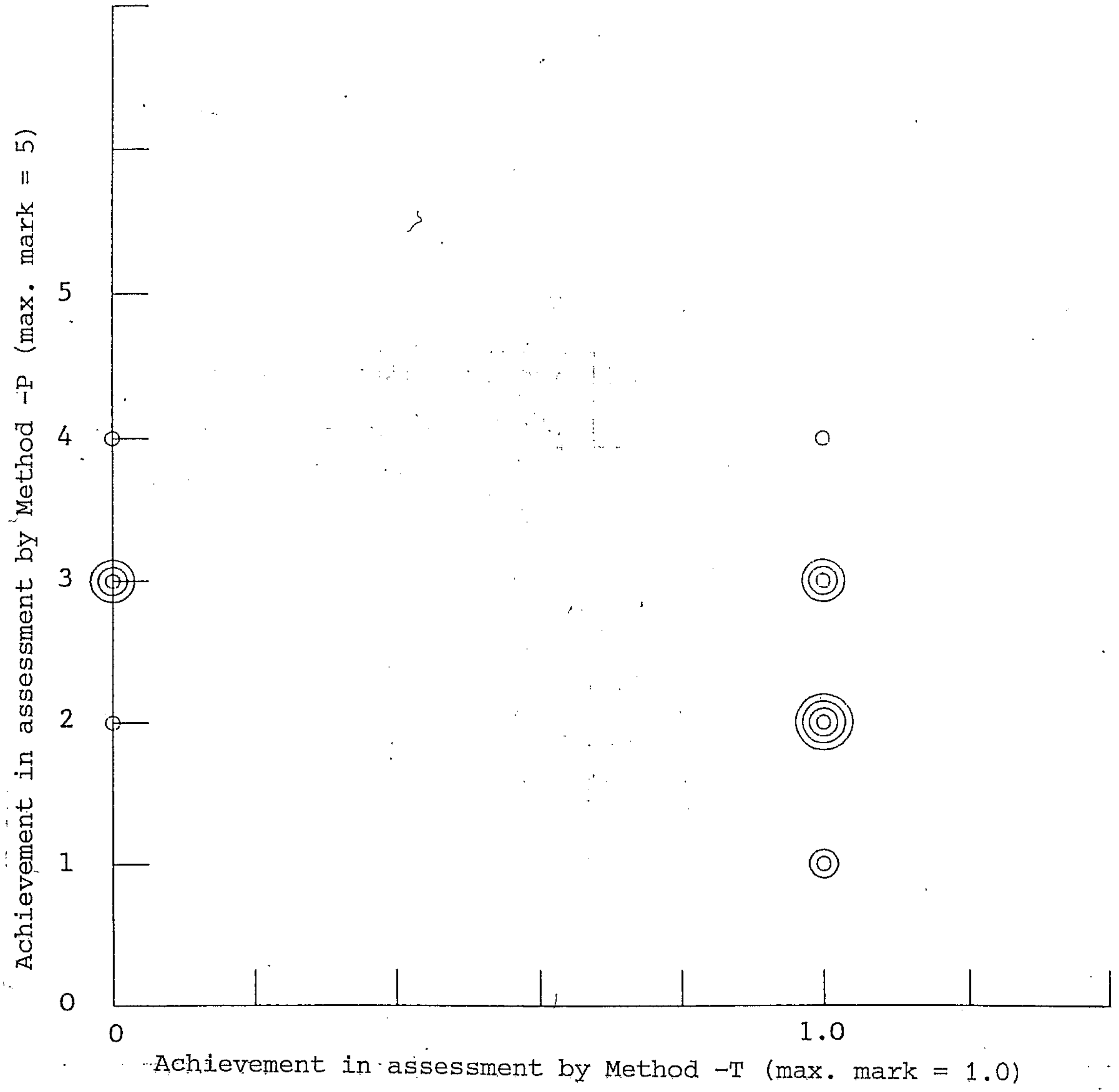

Fig. 4.2 Plots of the achievement data obtained from assessment by Method -T for measuring volume using burette by pupils of the "heating solids" group (those who got lessons on heating solids) participating in both the sessions against their achievement data obtained from assessment by Method -P for carrying out the whole experiment 4.3.

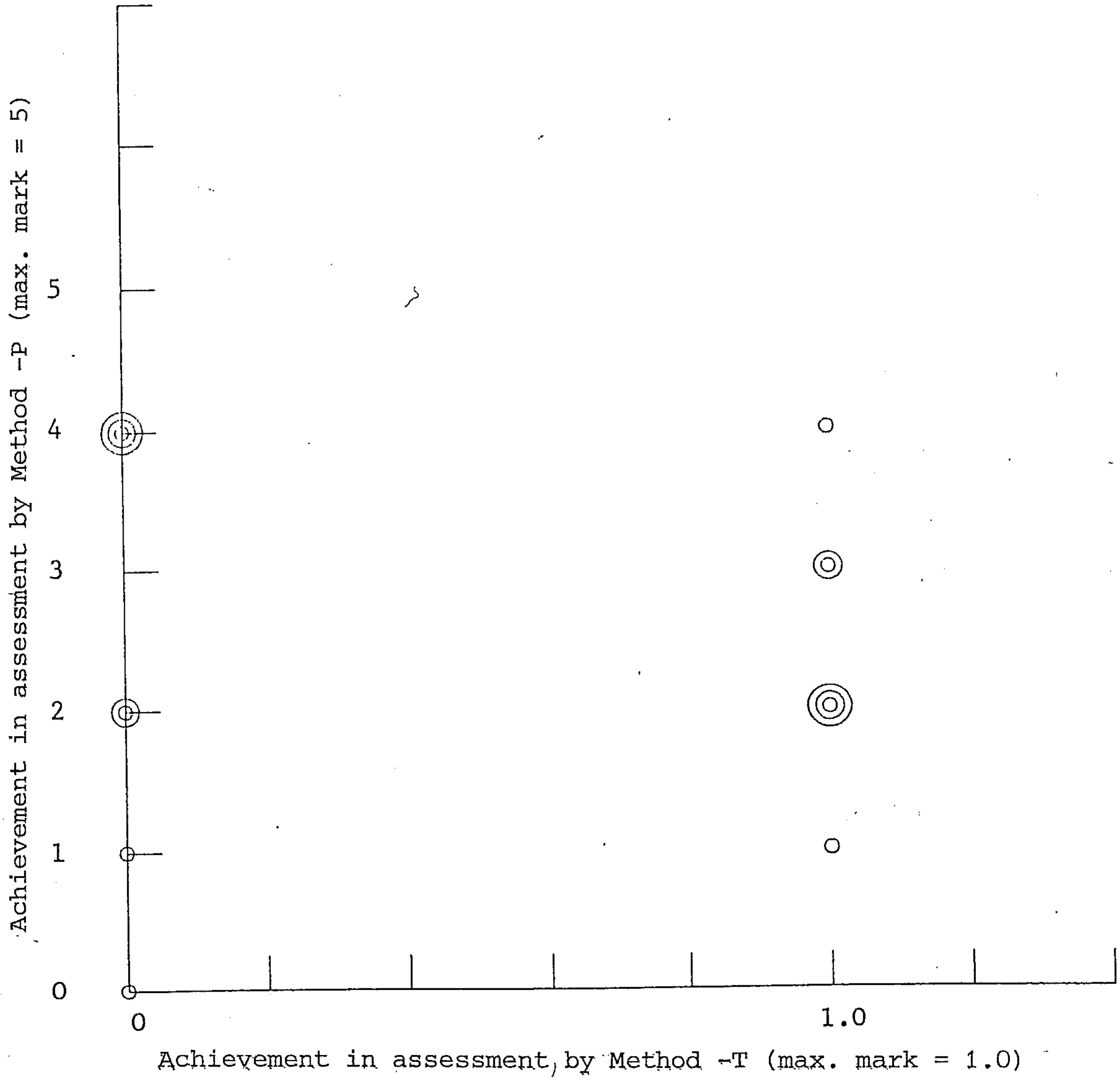

Fig. 4.3 Plots of the achievement data obtained from assessment by Method -T for measuring volume using burette by pupils of the "Control" group (those who did not get any lessons) participating in both the sessions against their achievement data obtained from assessment by Method -P for carrying out the whole experiment 4.3.